Received Date : 30-Apr-2016

Revised Date : 02-Aug-2016

Accepted Date : 04-Aug-2016

Article type : Original Article

1. Title:

Tests of species-specific models reveal the importance of drought in postglacial range shifts of a Mediterranean-climate tree: insights from iDDC modelling and ABC model selection

2. Authors:

Jordan B. Bemmels[1,3], Pascal O. Title[1], Joaquín Ortego[2], L. Lacey Knowles[1]

3. Postal addresses:

[1] Department of Ecology and Evolutionary Biology, 830 N. University Ave., University of Michigan, Ann Arbor, MI, USA 48109

[2] Department of Integrative Ecology, Estación Biológica de Doñana, EBD-CSIC, Avda. Américo Vespucio s/n, E-41092 Seville, Spain

[3] Corresponding author; email: jbemmels@umich.edu

4. Key words:

Approximate Bayesian Computation, California Floristic Province, climate change, drought, genetic structure, iDDC modelling

5. Contact information for corresponding author:

Jordan Bemmels
jbemmels@umich.edu
734-709-2914

University of Michigan
2010A Kraus Natural Science Building
830 N. University Ave.
Ann Arbor, MI 48109
USA

6. Running title

Drought and genetic structure of a live oak

ABSTRACT

Past climate change has caused shifts in species distributions and undoubtedly impacted patterns of genetic variation, but the biological processes mediating responses to climate change, and their genetic signatures, are often poorly understood. We test six species-specific biologically informed hypotheses about such processes in canyon live oak (*Quercus chrysolepis*) from the California Floristic Province. These hypotheses

encompass the potential roles of climatic niche, niche multidimensionality, physiological trade-offs in functional traits, and local-scale factors (microsites and local adaptation within ecoregions) in structuring genetic variation. Specifically, we use ecological niche models (ENMs) to construct temporally dynamic landscapes where the processes invoked by each hypothesis are reflected by differences in local habitat suitabilities. These landscapes are used to simulate expected patterns of genetic variation under each model and evaluate the fit of empirical data from 13 microsatellite loci genotyped in 226 individuals from across the species range. Using Approximate Bayesian Computation (ABC), we obtain very strong support for two statistically indistinguishable models: a trade-off model in which growth rate and drought tolerance drive habitat suitability and genetic structure, and a model based on the climatic niche estimated from a generic ENM, in which the variables found to make the most important contribution to the ENM have strong conceptual links to drought stress. The two most probable models for explaining patterns of genetic variation thus share a common component, highlighting the potential importance of seasonal drought in driving historical range shifts in a temperate tree from a Mediterranean climate where summer drought is common.

INTRODUCTION

Shifts in species distributions in response to climate change are a key factor structuring population genetic variation in both temperate and tropical species (Taberlet *et al.* 1998; Soltis *et al.* 2006; Carnaval *et al.* 2009; Morgan *et al.* 2011; Massatti & Knowles 2016). However, the biological mechanisms governing these shifts and their potential impact on patterns of neutral genetic variation are often poorly understood. For example, some plant species may be associated with ecological microsites partly or wholly defined by non-climatic factors (e.g., John *et al.* 2007; Frei *et al.* 2012; Allié *et al.* 2015) that could constrain responses to regional-scale climate change (Kroiss & HilleRisLambers 2015). Likewise, geographic distributions may be limited by different abiotic stresses (e.g. cold temperatures, drought) among species (Normand *et al.* 2009), or by different factors in different geographic regions of a single species' range (Morin *et*

al. 2007). Consequently, more detailed species-specific hypotheses about the causes of range shifts and their impacts on population genetic structure are needed (Papadopoulou & Knowles 2016). To this end, we develop and test a suite of competing biologically-informed models (Table 1) to explain the genetic structure of canyon live oak (*Quercus chrysolepis* Liebm., Fagaceae). These models make different predictions about patterns of genetic variation, depending upon the relative importance of climatic niche, niche multidimensionality, physiological trade-offs in functional traits, and local-scale factors (e.g., microsites and local adaptation within ecoregions) in governing the species' distribution and demographic history since the Last Glacial Maximum (LGM, 21.5 ka).

Considering that canyon live oak is a member of the climatically and ecologically heterogeneous California Floristic Province (CFP) of western North America and is distributed across a wide range of elevations (90 to 2,740 m; Thornburgh 1990), the response of this species to shifts in climate might be associated with different aspects of its ecology. For example, canyon live oak grows on many soil types and in many forest and chaparral communities (Thornburgh 1990), but is found exclusively in regions of high topographic complexity (Little 1971). Likewise, it is common throughout California, Oregon, and Baja California (Fig. 1), but is most abundant in sheltered canyons and on steep, rocky slopes, where it may be the dominant tree species (Thornburgh 1990). Consequently, while regions with climates similar to those of its present distribution likely existed in California's flat Central Valley during the LGM (Ortego *et al.* 2015), the climatic niche by itself may not accurately represent past distributional shifts in regions where topographic complexity is very low. Alternatively, it is possible that shifts in distributions due to past climate change might reflect constraints due to trade-offs in functional and physiological traits. For example, a trade-off between drought tolerance and growth rate may exist in species from climates with hot, dry summers (Howe *et al.* 2003; Alberto *et al.* 2013; Aitken & Bemmels 2016), and drought determines range limits of some plant species, including trees (Morin *et al.* 2007; Normand *et al.* 2009; Linares & Tíscar 2011; Rasztovits *et al.* 2014; Urli *et al.* 2014). Moreover, in many temperate trees, a trade-off between growth rate and cold tolerance drives population-level local

adaptation (Howe *et al.* 2003; Savolainen *et al.* 2007; Alberto *et al.* 2013; Aitken & Bemmels 2016) and may determine species range limits (Loehle 1998; but, see Morin *et al.* 2007 for a counterperspective). Given geographic variation in functional traits in many tree species, it is also possible that geographic range shifts in response to climate change will depend strongly on individual responses of specific populations to unique environmental factors (e.g., Davis *et al.* 2001; Pearman *et al.* 2010; Benito Garzón *et al.* 2011; Valladares *et al.* 2014; Gotelli & Stanton-Geddes 2015; Hällfors *et al.* in press). Lastly, the response to past climate change might simply reflect shifts in habitat suitability as it relates to basic climate variables, without the need to invoke complex, species-specific nuances of niche or mechanistic trade-offs in functional traits. Basic climate variables (e.g., temperature, precipitation) are frequently used in correlative ecological niche models (ENMs) to model species distributions and to predict how distributions have changed over time (Alvarado-Serrano & Knowles 2014). In canyon live oak specifically, previous work has shown that patterns of genetic connectivity and admixture among populations are correlated with areas of high habitat suitability since the LGM, as predicted by a climatic ENM (Ortego *et al.* 2015).

It is these types of biologically informed hypotheses that motivate this study (as opposed to generic statistical phylogeographic tests; reviewed in Papadopoulou & Knowles 2016). Specifically, through tests of six models (Table 1) we explore the relative support for alternative hypotheses about the niche of canyon live oak and factors that may have driven its response to climate change, including basic climate variables, microsites, niche multidimensionality, trade-offs in functional traits, and local adaptation within ecoregions. We use integrative distributional, demographic and coalescent (iDDC) modelling (Knowles & Alvarado-Serrano 2010; Brown & Knowles 2012; He *et al.* 2013) to generate genetic expectations under each model, and Approximate Bayesian Computation (ABC; Beaumont *et al.* 2002; Csilléry *et al.* 2010) to evaluate the fit of empirical data characterized from 13 microsatellite loci in 226 indiviuals sampled across the species range to the genetic predictions of each model. We highlight how careful extraction of spatially explicit information from ENMs reflecting the different processes that may influence range shifts in response to past climate

change is a key step in translating biologically-informed species-specific hypotheses into testable genetic predictions about a species' response to climate change.

MATERIALS AND METHODS

Sampling and genotyping

We collected leaf tissue from a total of 257 adult individuals from 46 localities across California (Fig. 1; Table S2); 160 individuals were sampled by Ortego *et al.* (2015), and 97 additional individuals were collected to provide complete geographic sampling for this study. Samples were genotyped at 13 polymorphic nuclear microsatellite markers developed for use in *Quercus* (Steinkellner *et al.* 1997; Kampfer *et al.* 1998; Durand *et al.* 2010). Full characterization of microsatellite loci and DNA extraction and microsatellite genotyping followed the procedures described by Ortego *et al.* (2014, 2015). Only individuals that were successfully genotyped at 10 or more of the 13 loci were retained for subsequent analyses (see Table S2), resulting in a dataset with a total of 226 individuals from 44 localities.

Assignment of individuals into populations

Populations were initially classified geographically based on major mountain ranges. Individuals were also assigned to different genetic clusters on the basis of their microsatellite genotypes using the Bayesian analysis implemented in *STRUCTURE v.2.3.4* (Pritchard *et al.* 2000; Falush *et al.* 2003, 2007; Hubisz *et al.* 2009). The likelihood of different genetic clusters (K = 1 to 10) was estimated from 10 independent runs with one million MCMC cycles, following a burn-in step of 100,000 iterations. *STRUCTURE* was run both with and without a prior conditioned on either individual sampling localities or the mountain ranges of sampled localities (Hubisz *et al.* 2009). Genetic clusters generally corresponded well to mountain ranges, except for localities from the Sierra Nevada. Sierra Nevada localities were often assigned to two different genetic clusters – a group of northern and of southern localities (Fig. S1). As a result of these analyses, we divided the 226 individuals from 44 localities into six populations, which included the Peninsular

Ranges, Transverse Ranges, Southern Sierra Nevada, Northern Sierra Nevada, Southern Coast Ranges, and Northern Coast Ranges and Klamath Mountains (Fig 1, Table S2). A Mojave Desert population was excluded from all further analyses due to small sample size (n = 6).

Translating hypotheses into ecological niche models

Ecological niche models (ENMs) were used to generate habitat suitability maps for canyon live oak in the present and during the Last Glacial Maximum (LGM, 21.5 ka), using maximum entropy modelling with *Maxent v.3.3.3k* (Phillips *et al.* 2004, 2006). Details of the general niche modelling procedure and data sources are given in the Supporting Information. To construct ENMs, specific environmental variables were selected as proxies for the biological mechanisms hypothesized to determine habitat suitability, as summarized below (see Table S1 for complete details of all variables included in each model, and the *Supplemental Methods* for more detailed justification of variable selection):

(1) *GeneralENM*: This model does not invoke a specific mechanism determining geographic range, but focuses on the assumption that basic climatic variables (Table S1; Hijmans *et al.* 2005) characterize habitat suitability according to a generic climatic ENM.

(2) *Microsite*: This model focuses on the assumption that habitat suitability may be limited by the availability of specific microsites such as canyons, steep slopes, and mountain ridges where canyon live oak could have a competitive advantage over other tree species (Thornburgh 1990). We assume that the four topographic variables that are included in this model (elevation, slope, aspect, and terrain roughness index; Amante & Eakins 2009; Hijmans *et al.* 2015a; Title & Bemmels in prep) have not substantially changed within the CFP since the LGM, except for exposed continental shelf due to lower sea levels and increased extent of glaciation during the LGM (see Supporting Information).

(3) *Multidimension*: This model assumes that a combination of basic climate variables, microsite, and additional climate variables putatively more closely related to

ecological processes (Table S1; Wang *et al.* 2006, 2012; Golicher 2012; Metzger *et al.* 2013; Title & Bemmels in prep) determines habitat suitability. These variables include all variables from the *GeneralENM* and *Microsite* models (but excluding elevation), as well as additional ecologically-relevant variables summarizing evapotranspiration, thermicity, aridity, growing degree days, and length of the growing season (Table S1). Note that elevation was excluded because the relationship between elevation and climate under current conditions is very different from the relationship that existed during the LGM (Ritter & Hatoff 1975).

(4) *GrowCold:* This model focuses on a possible trade-off between growth rate and cold tolerance that may constrain suitable habitat of canyon live oak. The model is constructed from variables hypothesized to reflect the level of abiotic stress and selective pressure experienced by the species and its fitness relative to competitors in relation to this trade-off (Table S1). We include variables related to cold-induced stress (e.g., mean temperature of the coldest quarter) as well as ameliorating variables indicating opportunity for growth during non-stressful conditions (e.g., growing degree days ≥ 5ºC).

(5) *GrowDrought*: This model focuses on a possible trade-off between growth and drought tolerance that may constrain suitable habitat of canyon live oak. As in the *GrowCold* model, chosen variables are hypothesized to reflect the level of abiotic stress experienced by the species and potential impacts on its fitness relative to competitors in relation to this trade-off (see Table S1); both stressor and ameliorating variables were included (as discussed above).

(6) *LocalAdaptation*: As in the *Multidimension* model, all available climatic and topographic variables (except elevation) are used to construct the ENM for this model, but with the difference that populations within each region are hypothesized to be strongly locally adapted. As such, habitat suitability in this model is predicted by unique climatic and topographic variables for each region separately, rather than the species as a whole (see also Gray & Hamann 2013). Given that genetic expectations are generated for the entire species range (as detailed below), regional habitat-suitability maps were standardized and combined into a single map (i.e., the habitat-suitability value of each

grid cell in the combined map was set equal to the highest habitat suitability for the corresponding grid cell in any of the individual regional maps). Regions of local adaptation were delimited using Commission for Environmental Cooperation North American Level III Ecoregions (CEC 1997), retaining only ecoregions with at least 25 occurrence records. A total of six ecoregions met this criterion: California Coastal Sage, Chaparral, and Oak Woodlands; Coast Range; Klamath Mountains; Mojave Basin and Range; Sierra Nevada; and Southern and Baja California Pine-Oak Mountains. Each ecoregion comprised an average of 231 occurrence records (range: 47 to 401). The ecoregion-based population definitions described here were used only for the purpose of constructing ENMs in the *LocalAdaptation* model. Note also that such localized effects of ecoregion-specific habitat suitabilities were only investigated with respect to the same bioclimatic variables as in the *Multidimension* model (and not with respect to the subsets of bioclimatic variables featured in each of the other four models) because of computational limitations.

Genetic predictions of each model

The integrative distributional, demographic and coalescent (iDDC) approach (He *et al.* 2013) was used to generate genetic predictions under each model (Fig. 2). For each separate model, (i) relative habitat suitabilities were extracted from the spatially explicit distributional model provided by the ENM, and were then rescaled to inform carrying capacities and migration rates of (ii) a demographic expansion across the landscape. For each of the six models tested, demographic simulations were conducted on landscapes representing three consecutive time periods, with corresponding shifts in habitat-suitability in response to changes in climate since the Last Glacial Maximum (LGM) for each time period. Specifically, maps for the present time period and for the LGM were generated directly from projections of the ENMs; a map representing intermediate conditions was also generated, in which the value of each grid cell corresponds to the mean value of that grid cell in the present and LGM maps. Parameters from the spatially explicit demographic model were then used to (iii) generate genetic predictions under a spatially explicit coalescent simulation. Finally,

datasets simulated under the iDDC models were compared with the empirical data using an Approximate Bayesian Computation (ABC) framework for model selection and parameter estimation (Beaumont *et al.* 2002).

Demographic simulations were conducted in *SPLATCHE2* (Ray *et al.* 2010) and were initiated at 21.5 ka from hypothesized ancestral source populations for each model. Ancestral source populations were defined as all grid cells of the LGM map with habitat suitability greater than the median habitat suitability of all grid cells of the current climate map containing an occurrence record (Brown & Knowles 2012). This threshold averaged 0.57 among models (range: 0.52 to 0.59). Note that relative LGM habitat suitability was obtained from each model directly as output of the ENM produced in *Maxent* (on a scale from 0 to 1). Next, habitat-suitability values for all maps across all time periods were categorized into 20 bins of equal magnitude, and maps were then used to perform the spatially explicit demographic simulations. In the demographic simulations, population carrying capacities and migration rates of each grid cell were rescaled proportionally according to habitat-suitability bins (with carrying capacity and migration rate ranging from zero to the maximum value of these parameters in a given simulation, as sampled from the prior distribution; see below). Note that because a single map is required by *SPLATCHE2,* custom *Python* scripts (provided by Q. He and deposited in *Dryad*, see *Data Accessibility* section) were used to convert the three maps of 20 bins each (39 bins for the intermediate map to account for intermediate values averaged between two bins; see above) into a single map with a theoretical maximum of 20^2x39 categories, with each category representing a unique combination of habitat-suitability bins across the three time periods. This makes it possible to model a dynamic landscape where habitat suitabilities change over time. Habitat-suitability bins representing each of the three temporal periods (LGM, intermediate, current) were consecutively applied for one third of the total number of generations each. Given that reproductive maturity in canyon live oak occurs after 15-20 years but individuals may live up to 300 years (Thornburgh 1990), average generation time was assumed to be 50 years, resulting in 430 generations from the LGM to present.

Following each time-forward demographic simulation, a time-backward coalescent genetic simulation was performed, in which the ancestry of an allele was traced back from the present into ancestral source populations. Before the the onset of population expansion from suitable areas at 21.5 ka modelled by the ENMs (see Fig. 3), alleles coalesced in a single large ancestral population (a maximum of 10^7 generations used in the simulations provided ample time for coalescence).

Individuals in simulated datasets were sampled from the same grid cells corresponding to the geographic locations from which the empirical data were sampled, and genetic data for these individuals were simulated along the coalescent genealogies at each locus using a strict stepwise microsatellite mutational model assuming no indels of more than one repeat unit, no recombination, and a maximum number of alleles equal to the number of repeat units separating the largest and smallest allele for each locus in the empirical data.

Model selection and parameter estimation using ABC

For the empirical data (Table S3) and each simulated genetic dataset, 24 summary statistics were calculated (mean, total, and population heterozygosity, *H*; total and population pairwise population differentiation, F_{ST}) using *Arlequin v.3.5* (Excoffier & Lischer 2010). Although the number of alleles, *K*, has previously been used as a summary statistic (He *et al.* 2013), it was not used here because *K* was difficult to fit to empirical data in simulations across all models (i.e., all models had a consistent tendency to generate values of *K* substantially lower than in the empirical data; see Table S4). We were thus concerned that the distance threshold between empirical and simulated datasets would need to be very large in order to retain a sufficient number of simulations for parameter estimation, which may have reduced the precision of parameter estimates (Beaumont *et al.* 2002). To check whether excluding *K* would have a major impact on model selection, we conducted simulations to validate our model selection procedure (validation methods described below) with and without *K*, and found that including *K* had very little impact on our ability to distinguish among models (results not shown). We also note that our models are highly capable of producing datasets with

properties that match the empirical data with respect to the 24 summary statistics used here (Table 1 and S4).

Rather than estimating parameter posterior distributions directly from summary statistics, partial least squares (PLS) components were calculated from summary statistics in order to reduce the number of summary statistics and account for correlations among them (Boulesteix & Strimmer 2006) using the *transformer* tool in *ABCtoolbox* with boxcox transformation for the pooled first 10,000 runs of each model (following He *et al.* 2013). In order to determine the optimal number of PLS components to retain, root mean squared error (RMSE) plots were examined and five PLS components were retained for calculating the distance between simulations and the empirical observations, because RMSE of the four parameters in our models does not decrease substantially with additional PLS components (results not shown).

Approximate Bayesian Computation (ABC) was used to estimate parameters and select among our six models using the wrapper program *ABCtoolbox* (Wegmann *et al.* 2010) on a high-performance computing cluster (Advanced Research Computing at the University of Michigan). One million datasets were simulated for each model across a broad range of parameter values (i.e., maximum carrying capacity, K_{max}; migration rate, m; ancestral effective population size before population expansion, N_{anc}; and microsatellite mutation rate, μ) under a uniform prior on the base 10 logarithm of each parameter. The priors for parameter values were the same among models (i.e., $log(K_{max})$, 2.7 to 4.0; $log(N_{anc})$, 3.0 to 6.0; $log(\mu)$, -6.0 to -2.0; and $log(m)$, -3.0 to -1.7; Fig. 4), with the exception of the *GrowCold* model, for which higher values of $log(m)$ were used ($log(m)$, -2.6 to -1.3) to ensure colonization of interior areas (see below). Note that the *GrowCold* model was the only model for which exclusively coastal ancestral source populations were inferred (Fig. S2). Because the same range of parameter values was used in all models, this different prior in the *GrowCold* model is unlikely to have biased model selection given that the density of simulations for the given range of parameter space was the same in all models.

In all models, priors on migration rate were carefully considered in order to reflect (a) biologically realistic values of migration rate, and (b) values that would result in

colonization of the landscape within the time spanning the LGM to the present. For example, true migration rates of our species are not known, but the prior -3.0 ≤ *log(m)* ≤ -1.7 covers potentially high values of migration rates at the spatial and temporal scale of our simulations (5-arcminute or ~9km x 9km grid cells; 50 years per generation) and we tested a variety of migration rates (and carrying capacities) in initial simulations to identify a range of migration rates that would result in colonization of the landscape within the time spanning the LGM to the present. Specifically, we identified a minimum value of *log(m)* for which complete landscape colonization was achieved (i.e., lower values were not included in the prior for *log(m)* because the landscape would not be completely colonized, which could bias model selection). Likewise, we did not apply exceptionally high *log(m)* values because such values resulted in such rapid colonization that the differences among models in terms of their colonization patterns would be lost.

For each model, 5,000 simulations (0.5% of the total number of simulations per model) that most closely matched those of the empirical data were retained (He *et al.* 2013) and used to generate posterior distributions of parameters, using ABC-GLM (general linear model) adjustment (Leuenberger & Wegmann 2010). Bayes factors were approximated in order to assess relative support for the most strongly supported model compared to each other model; the approximate Bayes factor in favour of model X over model Y is calculated as the marginal density of model X divided by the marginal density of model Y (Leuenberger & Wegmann 2010).

Validation of model choice and parameter estimates

To determine whether the alternative models can be accurately distinguished with ABC given the data, we simulated 100 pseudo-observed datasets (PODs) under each model and analyzed them using our ABC procedure for model choice, using a subset of total simulations (100,000 per model) for computational efficiency. For each model, we calculated the proportion of the PODs for which the true model was either correctly or incorrectly identified. For PODs for which the true model was correctly chosen, the strength of support for the true model was calculated as the mean logarithm of the

Bayes factor comparing the true model to the model with the second-highest marginal density. This represents how strongly the true model is identified to the exclusion of all other models. When an incorrect model was chosen, the strength of support for the incorrect model was calculated as the mean logarithm of the Bayes factor comparing the incorrectly chosen model to the true model used to generate the POD. This value determines how strongly the incorrect model is favoured over the true model. Lastly, to assess the ability of each model to generate the empirical data, Wegmann's *et al.* (2010) *p*-value was calculated from 5,000 retained simulations. This *p*-value is the proportion of simulated datasets with a smaller or equal likelihood than the empirical data under the ABC-GLM (Wegmann *et al.* 2010).

To assess the accuracy of parameters estimated with ABC, we calculated the posterior quantiles of true parameter values from 1,000 PODs for the models with highest support. A Kolmogorov-Smirnov test was used to test these quantiles against a uniform distribution. Deviation from a uniform distribution indicates bias in parameter estimation (Cook *et al.* 2006; Wegmann *et al.* 2010).

To determine whether there are specific summary statistics that are easier or more difficult to fit to the empirical data in specific models, we generated a distribution of the simulated values of each summary statistic from 100,000 simulations per model (with simulation parameters drawn from the prior). We then calculated the percentile corresponding to the empirical value of each summary statistic within its simulated distribution, and calculated the distance between this percentile and the median (i.e., 50th percentile) of the simulated distribution.

RESULTS

Multiple disjunct putative ancestral source populations based on habitat suitability during the LGM were estimated under each of the six models (Figs. 3 and S2). These sources included locations in both coastal and inland mountain ranges, with the exception of exclusively coastal ancestral source populations estimated for the *GrowCold* model. Predicted habitat suitability during the LGM and intermediate time

periods differed substantially among the six models, with the exception of the *GeneralENM* and *GrowDrought* models, which had very similar predictions for these time periods. In contrast, the current distribution of predicted suitable habitat was very similar for all models, except that the *Microsite* model also predicted large areas outside of the species' current range to contain suitable habitat (Figure S3).

With respect to the relative probabilities of the six models, two models – the *GeneralENM* model and the *GrowDrought* model – had the highest support (highest marginal density; Table 1). However, the Bayes factor comparing these two models was less than three, suggesting that there is not a statistically significant difference in the support for one model over the other (Kass & Raftery 1995). In other words, the *GeneralENM* and *GrowDrought* models are approximately equally well supported, in contrast to the much lower support for all the other models (Table 1). These two most probable models are also highly capable of generating simulated data comparable with the empirical data (see *p*-values, Table 1), despite uncertainty in parameter estimates (Fig. 4). Even with fairly broad posterior distributions for some parameter estimates (Fig. 4), the data contain information relevant to estimating the parameters (i.e., the posterior distribution differs from the prior), and there is evidence of increased accuracy of parameter estimates following GLM (general linear model) adjustment (Fig. 4). There is little evidence of bias in most parameter estimates (Fig. 5), except for slight deviations from uniformity detected from the quantiles of the mutation rate (μ) parameter for the *GeneralENM* and possibly the *GrowDrought* models (p = 0.0243 and 0.0503, respectively), and of the ancestral population size (N_{anc}) parameter for the *GrowDrought* model (p = 0.0082). A slight tendency to potentially overestimate each of these parameter values was detected (Fig. 5).

Validation of model selection using pseudo-observed datasets (PODs) showed that for most models, the true model is correctly identified the majority of the time (Table 2a) and average relative support for the true model is strong to very strong (Table 2b; Kass & Raftery 1995). Selection of an incorrect model with strong relative support is extremely uncommon. In the rare cases when an incorrect model is inferred, average relative support for the incorrectly chosen model compared to the true model is typically

very low (Table 2c), indicating that even if an incorrect model is identified as most likely, support is not strong enough to decisively exclude the true model from consideration. In contrast, for the *GeneralENM* and *GrowDrought* models there is limited ability to discern under which of these two models the PODs were simulated (Table 2). This is not surprising, given the similar relative support for these models in the empirical data (Table 1). Nonetheless, the *GeneralENM* and *GrowDrought* models are extremely unlikely to be confused with any of the other four models (Table 2).

Most models generated values of mean and total heterozygosity in agreement with empirical data, but simulated values of overall F_{ST} were typically higher than those of the empirical data in the *Multidimension*, *GrowCold*, and *LocalAdaptation* models (Table S4). These models also tended to produce certain population-specific simulated heterozygosity values that were lower than in the empirical data, and simulated pairwise F_{ST} values that were higher than in the empirical data. In contrast, the *Microsite* model tended to produce simulated pairwise F_{ST} values that were substantially lower than in the empirical data for comparisons involving the Northern Sierra Nevada population (and to a lesser extent, the Northern Coast Ranges and Klamath Mountains population). Simulated pairwise F_{ST} values involving the Northern Sierra Nevada population also tended to be lower than empirical values in the two most-supported models (*GeneralENM* and *GrowDrought*), although most other summary statistics in these models were similar to the empirical data.

DISCUSSION

Considering the biologically informed hypotheses we focus upon in our study, our goal was to consider whether we could distinguish among possible processes that might determine habitat suitability for canyon live oak and consequently, how the species distribution has shifted in response to changing climatic conditions. Differences in relative support among the models (Table 1) not only demonstrate differences in how influential these processes have likely been, but also how drought in particular may mediate the response to climate change in canyon live oak. Specifically, strong relative

support based on ABC model selection for two statistically indistinguishable models (Table 1) suggests that either climatic variables predictive of the species distribution that are related to drought stress (*GeneralENM* model), or a physiological trade-off between growth rate and summer drought tolerance (*GrowDrought* model), or both (see Table S1), are primary determinants of habitat suitability. More generally, this shared component of the two most highly supported models highlights the potential importance of drought in driving historical range shifts in a temperate tree from the predominately Mediterranean climate of the California Floristic Province (CFP), a region characterized by summer drought. Below, we discuss how our work contributes to an expanding literature about the factors that limit species distributions based on work from other disciplines, and compare and contrast our results with knowledge of factors important to other tree species from less seasonally dry regions of the temperate zone. We also discuss the implications of our work for evaluating support for alternative hypotheses (e.g., cold tolerance, microsite variation, and local adaptation) using explicit predictions for patterns of genetic variation, and the general challenges of our approach and the limitations of such inferences (see also Papadopoulou & Knowles 2016; Massatti & Knowles 2016).

Drought tolerance as a determinant of distributional shifts and genetic structure

In the Mediterranean climate of the CFP, summer is the driest season (Hijmans *et al.* 2005), and plants must tolerate or avoid summer drought stress. As such, summer drought is likely an important environmental condition determining relative habitat suitability for plants, either directly through abiotic stress or indirectly through effects on relative fitness in relation to competitors. The high support for the *GeneralENM* and *GrowDrought* models demonstrates that summer drought may not only be a key determinant of habitat suitability, but it may also drive demographic responses to climate change that ultimately impact population genetic structure of canyon live oak. In both of these models, the climatic variables making the largest contribution to the ENMs are strongly related to summer drought stress, and to the ability of a plant to tolerate or avoid this stress (see Table S1). The *GeneralENM* model uses a generic ENM in which

drought was not explicitly modelled and in which other climatic variables unrelated to drought were considered, but the four climatic variables making the greatest contribution to the ENM reflect precipitation during the summer and winter, and precipitation and temperature seasonality. As such, they represent the degree to which summers are hot and dry, and winters are cool and wet. Summer conditions likely directly reflect drought stress, whereas these winter conditions are hypothesized to reflect soil moisture availability during early spring, which may be the period of maximum growth for trees from Mediterranean environments prior to the onset of summer drought (Montserrat-Martí *et al.* 2009; Pinto *et al.* 2011). In comparison, the *GrowDrought* model features an ENM using climatic variables explicitly selected to reflect a possible trade-off between growth rate and summer drought tolerance. The climatic variables contributing most strongly to this ENM (Table S1) are preciptiation of the driest quarter and Emberger's pluviothermic quotient, which captures annual climatic dryness as experienced by plants with particular relevance to Mediterranean climates (Daget 1977).

The shared component of the two most supported models (i.e., drought stress) complements knowledge from other fields suggesting that drought limits geographic distributions and drives adaptation of some temperate tree species, especially those from Mediterranean climates. For example, across 1,577 European plant species, summer drought determines latitudinal range limits in 22% of species (Normand *et al.* 2009). Although drought stress does not generally limit the ranges of most of these plant taxa, its role in structuring plant distributions is especially common in the Mediterranean biomes of southern Europe, and in central Europe at the transition between Mediterranean and less seasonally dry biomes (Normand *et al.* 2009). Plant taxa with distributions limited by drought include trees specifically; for example, among European trees drought stress has been implicated in determining dry-edge range limits of *Fagus sylvatica* (Rasztovits *et al.* 2014), *Pinus nigra* (Linares & Tíscar 2011), and *Quercus robur* (Urli *et al.* 2014). Drought mortality was also found to be regionally important (e.g., in the Great Plains and at high-elevation sites) in limiting the ranges of at least 12 North America tree species (out of 17 studied; Morin *et al.* 2007).

In addition to setting range limits, drought tolerance is a trait of adaptive significance among populations of some tree species. For example, a trade-off between growth rate and drought tolerance has been documented among populations of Douglas-fir (*Pseudotsuga menziesii*; White 1987), and is hypothesized to underlie several adaptive differences in functional traits such as growth rate, growth phenology, growth pattern (i.e., determinate versus indeterminate), and root to shoot ratio (White 1987; Joly *et al.* 1989; Kaya *et al.* 1994). Putatively adaptive clines in phenotypic traits along precipitation gradients have also been observed in height growth and timing of bud flush in several western North American tree species (Aitken & Bemmels 2016). Although weak or non-adaptive clines along precipitation gradients may emerge when strong adaptive clines along temperature gradients exist (see below) if precipitation and temperature are geographically correlated, it is noteworthy that clines associated with precipitation are substantially stronger than those associated with temperature gradients in several species (e.g., *Picea pungens*, *Pinus attenuata, Pinus monticola, Populus trichocrapha*, and possibly *Pseudotsuga menziesii* and *Quercus garryana;* Aitken & Bemmels 2016).

While our procedure identified seasonal drought tolerance as an ecological factor that has likely shaped the response of canyon live oak to climate change and left signatures in patterns of genetic variation, our approach considers only the historically most important factors structuring genetic variation since the LGM. We tested only dynamic models (i.e., models where habitat suitability changes over time) because we have strong reason to believe that accounting for demographic history will be required to fully explain genetic structure in this study system. In particular, canyon live oak has a long generation time (we assumed only 430 generations since the LGM) and limited seed-dispersal ability by acorns (Thornburgh 1990), such that genetic signatures of past range shifts in response to climate change are unlikely to have been completely erased by contemporary patterns of gene flow (see Ortego *et al.* 2015). It is possible that ecological factors other than drought tolerance may be more important in driving contemporary processes affecting gene flow among populations, but testing these processes under contemporary climatic conditions was beyond the scope of our models.

Lack of support for competing explanations for genetic structure

Patterns of genetic variation in canyon live oak did not identify several commonly invoked competing factors (including cold tolerance, microsite variation, and local adaptation) as primary determinants of shifting geographic distributions in the face of climate change (Table 1). It is possible that this finding reflects differences in which environmental factors (e.g., temperature versus precipitation) are important for determining distributions and driving adaptation among different temperate tree species (see Howe *et al.* 2003; Normand *et al.* 2009; Aitken & Bemmels 2016). Yet, the lack of support for some of the models is nonetheless somewhat surprising, especially given that these models consider alternative ecological processes that are generally recognized to be broadly relevant across many taxa. For example, temperature is widely believed to limit cold-edge distributions in temperate trees through various physiological mechanisms (Sakai & Weiser 1973; Pigott & Huntley 1981; Morin *et al.* 2007; Normand *et al.* 2009; Mellert *et al.* 2011; Kollas *et al.* 2014; Lenz *et al.* 2014; Siefert *et al.* 2015). Furthermore, numerous tree species exhibit a trade-off between growth rate and cold tolerance at the population level, with more cold-tolerant populations exhibiting slower growth rate, earlier bud set, and (less frequently) shifts in phenology of bud flush (Howe *et al.* 2003; Savolainen *et al.* 2007; Alberto *et al.* 2013; Aitken & Bemmels 2016). This trade-off may also determine range limits at the species level, with warm-edge distributions limited by competition from faster-growing species and cold-edge distributions limited by low temperatures (Loehle 1998; but see also Morin *et al.* 2007). However, it is possible that the adaptive and ecological significance of drought in temperate trees has been understudied relative to that of cold temperatures because of biases in the choice of taxa studied. For example, most of the taxa studied are from temperate deciduous and conifer forests (Howe *et al.* 2003; Morin *et al.* 2007; Savolainen *et al.* 2007; Normand *et al.* 2009; Aitken & Bemmels 2016), whereas less attention has been paid to taxa from more seasonally dry regions of the temperate zone such as Mediterranean climates (e.g., Morin *et al.* 2007; Aitken & Bemmels 2016). In temperate broadleaf forests in particular, seasonal summer drought is uncommon and is

unlikely to be a major source of abiotic stress (Morin *et al.* 2007). The response to seasonal drought may also differ across biomes (Allen *et al.* 2010; Vicente-Serrano *et al.* 2013). In other words, temperate trees from Mediterranean climates may simply be subject to fundamentally different primary ecological and adaptive constraints than those from wetter, colder, and less seasonally dry climates within the temperate zone.

Lack of support for models reflecting alternative processes that could possibly affect habitat suitability (Table 1), especially those associated with local conditions, does not necessarily mean these processes do not play a role in response to climate change, but perhaps that their effects are minor at the regional scale studied here. In particular, lack of support for models incorporating local-scale factors (i.e., *Microsite* and *LocalAdaptation* models) suggests that responses to Pleistocene glacial cycles were primarily driven by climatic factors affecting habitat suitability over broad spatial scales. Consequently, although under current climatic conditions canyon live oak is distributed primarily in mountainous areas (Little 1971; Thornburgh 1990) and terrain roughness index (TRI) is one of the variables most highly predictive of current habitat suitability (*Multidimension* model; Table S1), TRI covaries with other predictor variables and may not itself be the driver of the species distribution. This interpretation also seems likely considering that both the *GeneralENM* and *GrowDrought* models receive high support, even though under these models the species is predicted to have been distributed in areas of low topographic complexity in the past (e.g., in California's northern Central Valley; Fig. 2). Our results are therefore consistent with the hypothesis that canyon live oak, despite its abundance in sheltered canyons and on steep, rocky slopes, was capable of making shifts to topographically novel habitats such as the northern Central Valley during the LGM (Fig. 2), which may reflect the ability of this species to grow on a wide variety of soil types and in multiple community assemblages (Thornburgh 1990).

Likewise, lack of support for the *LocalAdaptation* model (Table S1) suggests that the response of canyon live oak to climate change is not localized. Given that populations of many temperate and boreal tree species are locally adapted to climate (Savolainen *et al.* 2007; Alberto *et al.* 2013; Aitken & Bemmels 2016), local adaptation has been hypothesized to have been an important factor affecting Pleistocene range

623 shifts in trees (Davis *et al.* 2001), and is often considered to be a key factor that will
624 determine the effects of future climate change on the potential geographic distributions
625 of tree populations (e.g., Pearman *et al.* 2010; Benito Garzón *et al.* 2011; Gray &
626 Hamann 2013; Valladares *et al.* 2014; Gotelli & Stanton-Geddes 2015; Hällfors *et al.* in
627 press) and of adaptive genomic variation (Fitzpatrick & Keller 2015). In some cases,
628 local adaptation may also leave a signature in patterns of neutral genetic variation
629 (through its mediating effects on patterns of gene flow; e.g., Lee & Mitchell-Olds 2011).
630 While the *LocalAdaptation* model was not the most probable model identified in our
631 study, we note that it did receive very strong relative support compared to the
632 *Multidimension* model (Bayes factors = 234; Table 1) in which exactly the same
633 environmental variables were used to generate species-wide predictions of habitat
634 suitability (Table S1). This suggests that further investigation into localized effects of
635 other predictors of habitat suitability may indeed be worthwhile, especially with regards
636 to the highly supported models identified here (Table 1).
637 In addition to identifying the most probable models and determining that these
638 models are indeed capable of generating the data (Table 1), we also compared the
639 simulated summary statistics under each model with the empirical data (Table S4) to
640 examine what made a model a poor fit. This revealed that the empirical data did not
641 match the low heterozygosity and high pairwise F_{ST} values for certain populations
642 predicted by the *Multidimension*, *GrowCold*, and *LocalAdaptation* models. This lack of fit
643 suggests the generally small, disjunct ancestral source populations and spatially
644 restricted LGM habitat suitability predicted by these models (Figures S2, 3) is not well
645 supported by the data. In contrast, in the *Microsite* model, relatively low pairwise F_{ST}
646 values in the simulated data compared with the empirical data, especially for
647 comparisons involving the two northernmost populations, suggest that large areas of
648 high habitat suitability predicted since the LGM in the northern portion of this species'
649 range in this model (Figures S2, 3) are not well supported. A qualitatively similar pattern
650 (but with a smaller observed differences between simulated and empirical data) was
651 observed in both of the most well supported models (*GeneralENM, GrowDrought*),
652 suggesting even the most probable models do not capture the complex history of the

northern Sierra Nevada populations (Table S4). Exploring whether the northern Sierra Nevada historically contained smaller, more demographically isolated populations than suggested by our current models (Figures S2, 3) could be a hypothesis to test in future studies.

The California Floristic Province during the late Pleistocene

The California Floristic Province (CFP) is a plant biodiversity hotspot (Myers *et al.* 2000; Lancaster & Kay 2013) characterized by high topographic, climatic, and ecological heterogeneity. The maintenance of high biodiversity within the CFP has been hypothesized in part to reflect long-term regional-scale climatic stability that kept extinction rates low even through periods of intense global climatic change (Lancaster & Kay 2013). LGM habitat-suitability predictions for canyon live oak from the two most supported models (in fact, from all models except the *GrowCold* model; Figs. 3 and S2) are in agreement with this hypothesis. Both the *GeneralENM* and *GrowDrought* models predict high habitat suitability in some portion of every major mountain range in the CFP currently inhabited by the species, with the exception of the Mojave Desert and the northernmost portion of the range in the Klamath Mountains. The possible existence of these areas of high habitat suitability since the LGM throughout geographically disparate regions of the CFP suggests that canyon live oak is unlikely to have gone locally extinct in most regions of its current geographic distribution, and that only modest range shifts were needed in most regions in order for the species to track changes in suitable habitat.

This scenario contrasts with the major continental-scale changes in climate in response to glacial cycles that characterized other temperate regions such as eastern North America and Europe (Taberlet *et al.* 1998; Soltis *et al.* 2006; Gavin *et al.* 2014). At smaller spatial scales, pronounced effects of climate change did occur within the CFP. For example, alpine glaciers in the Sierra Nevada expanded in size (Gillespie *et al.* 2004), and pollen records indicate local changes in species abundance and shifts in the distribution of vegetation types to lower elevations (Roosma 1958; Cole 1983; Litwin *et al.* 1999; Heusser *et al.* 2015; McGann 2015), by as much as 600 to 750 m in the western Sierra Nevada (Ritter & Hatoff 1975). Nevertheless, at a regional scale, steep

683 elevational gradients and the moderating effects of orographic precipitation may have
684 provided a 'climatic buffering' effect preventing extreme regional-scale fluctuations in
685 climate (Lancaster & Kay 2013). As a result, species from the CFP were likely able to
686 track geographic shifts in suitable climate by migrating over relatively short distances
687 (Davis *et al.* 2008; Lancaster & Kay 2013). For canyon live oak in particular, large
688 regions of moderately stable habitat during both glacial and interglacial periods may
689 have served as reservoirs of genetic diversity and driven patterns of genetic connectivity
690 and admixture among populations (Ortego *et al.* 2015).
691
692 *Utility of species-specific genetic predictions for testing hypotheses*
693 Because different processes can produce similar patterns of genetic variation,
694 phylogeographic studies rely upon model-based inferences in which expectations for
695 patterns of genetic variation under particular processes are specified. However, the
696 approach applied here differs from other model-based inferences (see Knowles 2009;
697 Hickerson et al. 2010). Specifically, biologically informed hypotheses about factors that
698 may determine how taxa respond to climate change are explicitly modelled here by
699 considering their predicted effects on the movement of species across a landscape. As
700 such, our work adds to the growing number of studies that use spatially explicit models
701 to capture how population dynamics (e.g., changes in population size and dispersal
702 probabilities) impact patterns of genetic variation (e.g., Neuenschwander *et al.* 2008; He
703 *et al.* 2013; Massatti & Knowles 2014).
704 A key aspect of our approach – the generation of species-specific predictions for
705 patterns of genetic variation given different factors that might determine the habitat
706 suitability of a species – is a novel application that differs fundamentally from other
707 approaches for using patterns of genetic variation to study the effects of climate change
708 on geographic distributions of taxa. In particular, our approach considers that the best
709 characterization of habitat suitability for taxa may not be one based on a typical ENM
710 analysis of bioclimatic variables, as generally assumed in studies that rely on measures
711 of habitat suitability to test hypotheses about the effects of climate change using genetic
712 data (e.g., Knowles 2009; Lanier *et al.* 2015) There are nonetheless caveats with our

approach that should be considered, especially regarding the use of different environmental variables as proxies for competing biological processes hypothesized to determine habitat suitability. Specifically, we do not have an explicit means of determining if these environmental variables truly capture the processes they are intended to represent. This limitation is not unique to our approach. Instead, it is a broader conceptual concern with any approach in which predictions from correlative ENMs are used because it is not possible to ascertain whether environmental variables determine distributions directly, or are correlated with some other variable that is actually the source of causation but was not incorporated into the ENM (Austin 2002). While mechanistic ENMs that directly model functional traits of species could provide information to avoid misleading inferences about causal variables (Kearney & Porter 2009), the detailed information required for such functional modelling is frequently not available, which contrasts with the broad applicability of the approach applied here.

There are additional aspects of our study that should be kept in mind, some of which are not specific to our study, but are general issues with model-based inference. Our study provides a robust evaluation of competing models for observed patterns of genetic variation, as we evaluate not only the relative probabilities of models, but also conduct validations of our approach (i.e., we determine that the models are capable of generating the data, and that there is sufficient power to accurately distinguish among models given the quantity of genetic data collected in our study). As such, we can make strong statements about which of the different models best fit the data. However, we acknowledge there may of course be additional factors not considered here that might contribute to patterns of genetic variation, and therefore, our approach does not identify the optimal model (nor does any model-based approach). Recognizing the limits of the inference space is important for avoiding possible misinterpretations of model-based approaches, but it does not discount the insights gained with respect to the study goals. Instead, our work demonstrates that with thoughtful consideration of the factors that might determine habitat suitability (including not only climatic variables, but also potential trade-offs in functional traits that may impact a taxon's ability to tolerate physiological stresses or compete, as well as localized effects related to microsite

743 variation and adaptive differences), such hypotheses can be translated into models for
744 studying which factors mediate the effects of climate change on species distributions.
745 Likewise, even though many assumptions are made in the procedures applied here (e.g.,
746 converting measures of habitat suitability into population demographic parameters; for
747 details see Brown and Knowles 2012), these assumptions are arguably not more
748 problematic than many assumptions implicitly made in other model-based approaches
749 (e.g., not considering the spatial mosaic of habitat suitabilities that impacts both local
750 population sizes and migration probabilities, despite the clear effects of such
751 heterogeneity on patterns of genetic variation; see Knowles & Alvarado-Serrano 2010).
752 Lastly, spatially explicit models, despite some of their limitations discussed above (see
753 also Massatti & Knowles 2016), provide a window into a diversity of questions that would
754 continue to go unexplored without their application.
755
756 *Conclusions*
757 We compare the relative statistical support for six different models concerning
758 distributional shifts in canyon live oak in response to climate change, each of which is
759 motivated by a different hypothesis about the mechanistic factors that may determine
760 habitat suitability. We obtain very strong relative statistical support for two models that
761 share a common conceptual link to summer drought, and show through validation of the
762 model-selection procedure that we can be highly confident in the fit of data under these
763 models, as well as in our ability to accurately discriminate among the different models.
764 We suggest that drought tolerance may not only be a critical factor determining habitat
765 suitability and mediating distributional shifts in response to climate change since the
766 LGM in canyon live oak, but its importance may be generalized to other plants.
767 Specifically, by comparison with studies of other temperate trees that have emphasized
768 other processes but where focal taxa have typically been from less seasonally dry
769 regions of the temperate zone, our work suggests that summer drought may play a key
770 adaptive and ecologically important role in other trees from Mediterranean climates.
771 Moreover, our approach demonstrates how different factors hypothesized to determine
772 habitat suitability may be tested by using spatially explicit information from ENMs to

generate specific patterns of genetic variation for testing biologically informed hypotheses about the effects of climate change on species distributions. As such, the models supported in our study are a general example of the type of biologically informed, species-specific hypotheses that contribute to our broader understanding of the importance of biotic factors in structuring genetic variation (reviewed in Papadopoulou & Knowles 2016).

ACKNOWLEDGEMENTS

The authors thank P.F. Gugger for assistance with sampling; Q. He, R. Massatti, and A. Papadopoulou for their help and guidance in performing the ABC analyses; Q. He for providing custom scripts for iDDC modelling; and three anonymous reviewers for their suggestions, which greatly improved the manuscript. Funding was provided for fieldwork and microsatellite genotyping by an internal EBD 'Microproyectos' grant to J.O., financed by the Spanish Ministry of Economy and Competitiveness through the Severo Ochoa Program for Centres of Excellence in R+D+I (SEV-2012-0262); for a high-performance computing cluster allocation for simulations by a Collegiate Professor Honorarium at the University of Michigan (L.L.K.); for support to J.B.B. by an NSF GRFP fellowship (DEB: 1256260) and a University of Michigan Department of Ecology and Evolutionary Biology Edwin H. Edwards Scholarship in Biology; and for support to J.O. by a Ramón y Cajal Fellowship (RYC-2013-12501). This research was also supported in part through computational resources and services provided by Advanced Research Computing at the University of Michigan, Ann Arbor.

REFERENCES

Aitken SN, Bemmels JB (2016) Time to get moving: assisted gene flow of forest trees. *Evolutionary Applications*, **9**, 271–290.

Alberto FJ, Aitken SN, Alía R *et al.* (2013) Potential for evolutionary responses to climate change - evidence from tree populations. *Global Change Biology*, **19**, 1645–

1661.

Allen CD, Macalady AK, Chenchouni H *et al.* (2010) A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. *Forest Ecology and Management*, **259**, 660–684.

Allié E, Pélissier R, Engel J *et al.* (2015) Pervasive local-scale tree-soil habitat association in a tropical forest community. *PLOS ONE*, **10**, 1–17.

Alvarado-Serrano DF, Knowles LL (2014) Ecological niche models in phylogeographic studies: applications, advances and precautions. *Molecular Ecology Resources*, **14**, 233–248.

Amante C, Eakins BW (2009) ETOPO1 1 arc-minute global relief model: procedures, data sources and analysis. NOAA Technical Memorandum NESDIS NGDC-24. *National Geophysical Data Center, NOAA*.

Austin, MP (2002) Spatial predictions of species distribution: an interface between ecological theory and statistical modelling. *Ecological Modelling*, **157**, 101-118.

Beaumont MA, Zhang W, Balding DJ (2002) Approximate Bayesian computation in population genetics. *Genetics*, **162**, 2025–2035.

Benito Garzón M, Alía R, Robson TM, Zavala MA (2011) Intra-specific variability and plasticity influence potential tree species distributions under climate change. *Global Ecology and Biogeography*, **20**, 766–778.

Boulesteix A-L, Strimmer K (2006) Partial least squares: a versatile tool for the analysis of high-dimensional genomic data. *Briefings in Bioinformatics*, **8**, 32–44.

Brown JL, Knowles LL (2012) Spatially explicit models of dynamic histories: examination of the genetic consequences of Pleistocene glaciation and recent climate change on the American Pika. *Molecular Ecology*, **21**, 3757–3775.

Carnaval AC, Hickerson MJ, Haddad CFB, Rodrigues MT, Moritz C (2009) Stability predicts genetic diversity in the Brazilian Atlantic forest hotspot. *Science*, **323**, 785–789.

CEC (1997) *Ecological regions of North America: toward a common perspective*. Commision for Environmental Cooperation, Montréal.

Cole K (1983) Late pleistocene vegetation of Kings Canyon, Sierra Nevada, California.

Quarternary Reserach, **19**, 117–129.
Cook SR, Gelman A, Rubin DB (2006) Validation of software for Bayesian models using posterior quantiles. *Journal of Computational and Graphical Statistics*, **15**, 675–692.
Csilléry K, Blum MGB, Gaggiotti OE, François O (2010) Approximate Bayesian Computation (ABC) in practice. *Trends in Ecology and Evolution*, **25**, 410–418.
Daget P (1977) Le bioclimat méditerranéen: analyse des formes climatiques par le système d'Emberger. *Vegetatio*, **34**, 87–103.
Davis EB, Koo MS, Conroy C, Patton JL, Moritz C (2008) The California Hotspots Project: identifying regions of rapid diversification of mammals. *Molecular Ecology*, **17**, 120–138.
Davis MB, Shaw RG (2001) Range shifts and adaptive responses to Quaternary climate change. *Science*, **292**, 673–679.
Durand J, Bodénès C, Chancerel E *et al.* (2010) A fast and cost-effective approach to develop and map EST-SSR markers: oak as a case study. *BMC Genomics*, **11**, 570.
Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564–567.
Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, **164**, 1567–1587.
Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes*, **7**, 574–578.
Fitzpatrick MC, Keller SR (2015) Ecological genomics meets community-level modelling of biodiversity: mapping the genomic landscape of current and future environmental adaptation. *Ecology Letters*, **18**, 1–16.
Frei ES, Scheepens JF, Stöklin J (2012) Dispersal and microsite limitation of a rare alpine plant. *Plant Ecology*, **213**, 395–406.
Gavin DG, Fitzpatrick MC, Gugger PF *et al.* (2014) Climate refugia: joint inference from fossil records, species distribution models and phylogeography. *New Phytologist*,

204, 37–54.

Gent PR, Danabasoglu G, Donner LJ *et al.* (2011) The Community Climate System Model version 4. *Journal of Climate*, **24**, 4973–4991.

Gillespie A, Porter S, Atwater BF (2004) *Quaternary glaciations - extent and chronology. Part II: North America. Additional CD-ROM*. Elsevier, San Diego, CA.

Golicher D (2012) Implementing a bucket model using WorldClim layers. https://rpubs.com/dgolicher/2964. *RPubs*.

Gotelli NJ, Stanton-Geddes J (2015) Climate change, genetic markers and species distribution modelling. *Journal of Biogeography*, **42**, 1577–1585.

Gray LK, Hamann A (2013) Tracking suitable habitat for tree populations under climate change in western North America. *Climatic Change*, **117**, 289–303.

Hällfors MH, Liao J, Dzurisin JDK *et al.* (in press) Addressing potential local adaptation in species distribution models: implications for conservation under climate change. *Ecological Applications*.

Harrison SP, Bartlein PJ, Brewer S *et al.* (2014) Climate model benchmarking with glacial and mid-Holocene climates. *Climate Dynamics*, **43**, 671–688.

He Q, Edwards DL, Knowles LL (2013) Integrative testing of how environments from the past to the present shape genetic structure across landscapes. *Evolution*, **67**, 3386–3402.

Heusser LE, Kirby ME, Nichols JE (2015) Pollen-based evidence of extreme drought during the last Glacial (32.6-9.0 ka) in coastal southern California. *Quaternary Science Reviews*, **126**, 242–253.

Hickerson, MJ, Carstens BC, Cavender-Bares J *et al.* (2010) Phylogeography's past, present, and future: 10 years after Avise, 2000. *Molecular Phylogenetics and Evolution*, **54**, 291-301.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.

Hijmans RJ, van Etten J, Cheng J *et al.* (2015a) Package "raster". https://cran.r-project.org/web/packages/raster/index.html.

Hijmans RJ, Phillips S, Leathwick J, Elith J (2015b) Package “dismo”. http://cran.r-project.org/web/packages/dismo/index.html.

Howe GT, Aitken SN, Neale DB *et al.* (2003) From genotype to phenotype: unraveling the complexities of cold adaptation in forest trees. *Canadian Journal of Botany*, **81**, 1247–1266.

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, **9**, 1322–1332.

John R, Dalling JW, Harms KE *et al.* (2007) Soil nutrients influence spatial distributions of tropical tree species. *PNAS*, **104**, 864–869.

Joly RJ, Adams WT, Stafford SG (1989) Phenological and morphological responses of mesic and dry site sources of coastal Douglas-fir to water deficit. *Forest Science*, **35**, 987–1005.

Kampfer S, Lexer C, Glössl J, Steinkellner H (1998) Characterization of $(GA)_n$ microsatellite loci from *Quercus robur*. *Hereditas*, **129**, 183–186.

Kass RE, Raftery AE (1995) Bayes factors. *Journal of the American Statistical Association*, **90**, 773–795.

Kaya Z, Adams WT, Campbell RK (1994) Adaptive significance of intermittent shoot growth in Douglas-fir seedlings. *Tree Physiology*, **14**, 1277–89.

Kearney, M, Porter W (2009) Mechanistic niche modelling: combining physiological and spatial data to predict species’ ranges. *Ecology Letters*, **12**, 334-350.

Knowles LL (2009) Statistical phylogeography. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 593–612.

Knowles LL, Alvarado-Serrano DF (2010) Exploring the population genetic consequences of the colonization process with spatio-temporally explicit models: insights from coupled ecological, demographic and genetic models in montane grasshoppers. *Molecular Ecology*, **19**, 3727–3745.

Kollas C, Körner C, Randin CF (2014) Spring frost and growing season length co-control the cold range limits of broad-leaved trees. *Journal of Biogeography*, **41**, 773–783.

Kroiss SJ, HilleRisLambers J (2015) Recruitment limitation of long-lived conifers:

Implications for climate change responses. *Ecology*, **96**, 1286–1297.
Lancaster LT, Kay KM (2013) Origin and diversification of the California flora: re-examining classic hypotheses with molecular phylogenies. *Evolution*, **67**, 1041–1054.
Lanier HC, Massatti R, He Q, Olson LE, Knowles LL (2015) Colonization from divergent ancestors: glaciation signatures on contemporary patterns of genomic variation in Collared Pikas (*Ochotona collaris*). *Molecular Ecology*, 3688–3705.
Lee C-R, Mitchell-Olds T (2011) Quantifying effects of environmental and geographical factors on patterns of genetic differentiation. *Molecular Ecology*, **20**, 4631–4642.
Lenz A, Vitasse Y, Hoch G, Körner C (2014) Growth and carbon relations of temperate deciduous tree species at their upper elevation range limit. *Journal of Ecology*, **102**, 1537–1548.
Leuenberger C, Wegmann D (2010) Bayesian computation and model selection without likelihoods. *Genetics*, **184**, 243–252.
Linares JC, Tíscar PA (2011) Buffered climate chang eeffects in a Mediterranean pine species: range limit implications from a tree-ring study. *Oecologia*, **167**, 847–859.
Little EL (1971) *Atlas of United States trees, volume 1. Conifers and important hardwoods. USDA Miscellaneous Publication 1146.* Washington, DC.
Litwin RJ, Smoot JP, Durika NJ, Smith GI (1999) Calibrating Late Quaternary terrestrial climate signals: radiometrically dated pollen evidence from the southern Sierra Nevada, USA. *Quaternary Science Reviews*, **18**, 1151–1171.
Loehle C (1998) Height growth rate tradeoffs northern determine and southern range limits for trees. *Journal of Biogeography*, **25**, 735–742.
Massatti R, Knowles LL (2016) Tests of species-specific predictions for genetic variation in a wet versus dry-adapted species show that microhabitat preference of alpine plants may mediate the effects of climate change. *Molecular Ecology, in press*.
Massatti R, Knowles LL (2014) Microhabitat differences impact phylogeographic concordance of codistributed species: genomic evidence in montane sedges (*Carex* L.) from the Rocky Mountains. *Evolution*, **68**, 2833–2846.
McGann M (2015) Late Quaternary pollen record from the central California continental

margin. *Quaternary International*, **387**, 46–57.
Mellert KH, Fensterer V, Küchenhoff H *et al.* (2011) Hypothesis-driven species distribution models for tree species in the Bavarian Alps. *Journal of Vegetation Science*, **22**, 635–646.
Metzger MJ, Bunce RGH, Jongman RHG *et al.* (2013) A high-resolution bioclimate map of the world: A unifying framework for global biodiversity research and monitoring. *Global Ecology and Biogeography*, **22**, 630–638.
Montserrat-Martí G, Camarero JJ, Palacio S *et al.* (2009) Summer-drought constrains the phenology and growth of two coexisting Mediterranean oaks with contrasting leaf habit: Implications for their persistence and reproduction. *Trees*, **23**, 787–799.
Morgan K, O'Loughlin SM, Chen B *et al.* (2011) Comparative phylogeography reveals a shared impact of pleistocene environmental change in shaping genetic diversity within nine *Anopheles* mosquito species across the Indo-Burma biodiversity hotspot. *Molecular Ecology*, **20**, 4533–4549.
Morin X, Augspurger C, Chuine I (2007) Process-based modeling of species' distributions: what limits temperate tree species' range boundaries? *Ecology*, **88**, 2280–2291.
Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature*, **403**, 853–858.
Neuenschwander S, Largiadèr CR, Ray N *et al.* (2008) Colonization history of the Swiss Rhine basin by the bullhead (*Cottus gobio*): inference under a Bayesian spatially explicit framework. *Molecular Ecology*, **17**, 757–772.
Normand S, Treier UA, Randin C *et al.* (2009) Importance of abiotic stress as a range-limit determinant for European plants: insights from species responses to climatic gradients. *Global Ecology and Biogeography*, **18**, 437–449.
Ortego J, Bonal R, Muñoz A, Aparicio JM (2014) Extensive pollen immigration and no evidence of disrupted mating patterns or reproduction in a highly fragmented holm oak stand. *Journal of Plant Ecology*, **7**, 384–395.
Ortego J, Gugger PF, Sork VL (2015) Climatically stable landscapes predict patterns of genetic structure and admixture in the Californian canyon live oak. *Journal of*

983 *Biogeography*, **42**, 328–338.
984 Papadopoulou A, Knowles LL (2016) A paradigm shift in comparative phylogeography
985 driven by trait-based hypotheses. *PNAS,* **113**, 8018-8024.
986 Pearman PB, D'Amen M, Graham CH, Thuiller W, Zimmermann NE (2010) Within-taxon
987 niche structure: niche conservatism, divergence and predicted effects of climate
988 change. *Ecography*, **33**, 990–1003.
989 Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species
990 geographic distributions. *Ecological Modelling*, **190**, 231–259.
991 Phillips SJ, Dudík M, Schapire RE (2004) A maximum entropy approach to species
992 distribution modeling. *Proceedings of the Twenty-First International Conference on*
993 *Machine Learning*, 655–662.
994 Pigott C, Huntley JP (1981) Factors controlling the distribution of *Tilia cordata* at the
995 northern limits of its geographical range. III. Nature and causes of seed sterility.
996 *New Phytologist*, **87**, 817–839.
997 Pinto CA, Henriques MO, Figueiredo JP *et al.* (2011) Phenology and growth dynamics in
998 Mediterranean evergreen oaks: effects of environmental conditions and water
999 relations. *Forest Ecology and Management*, **262**, 500–508.
1000 Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using
1001 multilocus genotype data. *Genetics*, **155**, 945–959.
1002 Rasztovits E, Berki I, Mátyás C *et al.* (2014) The incorporation of extreme drought
1003 events improves models for beech persistence at its distribution limit. *Annals of*
1004 *Forest Science*, **71**, 201–210.
1005 Ray N, Currat M, Foll M, Excoffier L (2010) SPLATCHE2: a spatially explicit simulation
1006 framework for complex demography, genetic admixture and recombination.
1007 *Bioinformatics*, **26**, 2993–2994.
1008 Ritter EW, Hatoff BW (1975) Late Pleistocene pollen and sediments: an analysis of a
1009 central California locality. *Texas Journal of Science*, **29**, 195–207.
1010 Roosma A (1958) A climatic record from Searles Lake, California. *Science*, **128**, 716.
1011 Sakai AA, Weiser CJ (1973) Freezing resistance of trees in North America with
1012 reference to tree regions. *Ecology*, **54**, 118–126.

1013 Savolainen O, Pyhäjärvi T, Knürr T (2007) Gene flow and local adaptation in trees.
1014 *Annual Review of Ecology, Evolution, and Systematics*, **38**, 595–619.
1015 Sayre R, Comer P, Warner H, Cress J (2009) *A new map of standardized terrestrial*
1016 *ecosystems of the conterminous United States: US Geological Survey Professional*
1017 *Paper 1768*. Reston, VA.
1018 Siefert A, Lesser MR, Fridley JD (2015) How do climate and dispersal traits limit ranges
1019 of tree species along latitudinal and elevational gradients? *Global Ecology and*
1020 *Biogeography*, **24**, 581–593.
1021 Soltis DE, Morris AB, McLachlan JS, Manos PS, Soltis PS (2006) Comparative
1022 phylogeography of unglaciated eastern North America. *Molecular Ecology*, **15**,
1023 4261–4293.
1024 Steinkellner H, Fluch S, Turetschek E *et al.* (1997) Identification and characterization of
1025 $(GA/CT)_n$-microsatellite loci from *Quercus petraea*. *Plant Molecular Biology*, **33**,
1026 1093–1096.
1027 Taberlet P, Fumagalli L, Wust-Saucy A-G, Cosson J-F (1998) Comparative
1028 phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, **7**,
1029 453–464.
1030 Thornburgh DA (1990) Canyon live oak. In: *Silvics of North America: 1. Conifers; 2.*
1031 *Hardwoods. US Department of Agriculture Handbook 654.* (eds Burns RM, Honkala
1032 BH). Washington, DC.
1033 Thornthwaite CW (1948) An approach toward a rational classification of climate.
1034 *Geographical Review*, **38**, 55–94.
1035 Title PO, Bemmels JB (in prep) ENVIREM: an expanded set of bioclimatic variables
1036 improves ecological niche modeling performance. *Methods in Ecology and*
1037 *Evolution*.
1038 Urli M, Lamy J-B, Sin F *et al.* (2014) The high vulnerability of *Quercus robur* to
1039 drought at its southern margin paves the way for *Quercus ilex*. *Plant Ecology*, **216**,
1040 177–187.
1041 Valladares F, Matesanz S, Guilhaumon F *et al.* (2014) The effects of phenotypic
1042 plasticity and local adaptation on forecasts of species range shifts under climate

change. *Ecology Letters*, **17**, 1351–1364.
Vicente-Serrano SM, Gouveia C, Camarero JJ *et al.* (2013) Response of vegetation to drought time-scales across global land biomes. *PNAS*, **110**, 52–57.
Wang T, Hamann A, Spittlehouse DL, Aitken SN (2006) Development of scale-free climate data for western Canada for use in resource management. *International Journal of Climatology*, **26**, 383–397.
Wang T, Hamann A, Spittlehouse DL, Murdock TQ (2012) ClimateWNA—high-resolution spatial climate data for western North America. *Journal of Applied Meteorology and Climatology*, **51**, 16–29.
Wegmann D, Leuenberger C, Neuenschwander S, Excoffier L (2010) ABCtoolbox: a versatile toolkit for approximate Bayesian computations. *BMC Bioinformatics*, **11**, 116.
White TL (1987) Drought tolerance of southwestern Oregon Douglas-fir. *Forest Science*, **33**, 283–293.
Wilson MFJ, O'Connell B, Brown C, Guinan JC, Grehan AJ (2007) Multiscale terrain analysis of multibeam bathymetry data for habitat mapping on the continental slope. *Marine Geodesy*, **30**, 3–35.
Zhao L, Xia J, Xu C-Y *et al.* (2013) Evapotranspiration estimation methods in hydrological models. *Journal of Geographical Sciences*, **23**, 359–369.
Zomer RJ, Trabucco A, Bossio DA, Verchot L V. (2008) Climate change mitigation: A spatial analysis of global land suitability for clean development mechanism afforestation and reforestation. *Agriculture, Ecosystems & Environment*, **126**, 67–80.
Zomer RJ, Trabucco A, Van Straaten O, Bossio DA (2006) *Carbon, Land and Water: A Global Analysis of the Hydrologic Dimensions of Climate Change Mitigation through Afforestation/Reforestation. International Water Management Institute Research Report 101.* Colombo, Sri Lanka.

DATA ACCESSIBILITY

Microsatellite genotypes, occurrence records for ecological niche modelling, and custom *Python* scripts for ABC modelling are stored and accessible through the *Dryad* data repository, doi:10.5061/dryad.5gv12

AUTHOR CONTRIBUTIONS

All authors designed the research project. J.O. collected and genotyped samples and performed *STRUCTURE* analyses. J.B.B. and P.O.T. implemented the iDDC modelling and ABC procedures, under guidance of L.L.K. J.B.B. and L.L.K. wrote the manuscript, with suggestions from P.O.T. and J.O.

Table 1. Summary of models and relative support from the ABC procedure for each model. A higher marginal density corresponds to higher support for the model, while *p*-values close to 1.0 indicate that the model is able to reproduce data in agreement with the empirical data (Wegmann *et al.* 2010). Bayes factors represent the degree of relative support for the most highly supported model (*GeneralENM*) over the other models. Bayes factors greater than 20 indicate strong support, while those greater than 150 indicate very strong support (Kass and Rafterty 1995).

Model	Hypothesized factors mediating species response to climate change	Marginal density	Wegmann's *p*-value	Bayes factor
GeneralENM	basic climatic variables of a generic ecological niche model	2.35×10^{-2}	0.9900	-
Microsite	availability of topographic microsites	1.27×10^{-7}	0.0024	1.86×10^{5}
Multidimension	basic and ecologically-informed climate variables; microsites	8.20×10^{-9}	0.0038	2.87×10^{6}
GrowCold	trade-off between growth rate and cold tolerance	3.21×10^{-7}	0.0046	7.34×10^{4}
GrowDrought	trade-off between growth rate and drought tolerance	8.43×10^{-3}	0.9272	2.79
LocalAdaptation	unique factors in each locally adapted ecoregion	3.51×10^{-7}	0.0044	6.70×10^{4}

Table 2. Validation of the ABC procedure for model selection using pseudo-observed datasets (PODs; see text for explanation). (a) Confusion matrix showing the ability of the ABC procedure to correctly identify the model used to generate the POD. Numbers in the table represent the percent of PODs (n = 100 for each model) determined by the ABC procedure to be most highly supported by each of the models. Bold numbers on the diagonal indicate that the true model was identified, while numbers off the diagonal indicate incorrect model identification. (b-c) Average level of support, measured as the mean logarithm of Bayes factors, $log_{10}(BF)$, for (b) the true model compared to the second-most-supported model, when the true model is chosen, and (c) the incorrectly chosen model compared to the true model, when an incorrect model is chosen. Values in (b) represent the strength with which the ABC procedure unambiguously supports the true model to the exclusion of all other models, when the true model is chosen. Values in (c) represent the average strength with which the ABC procedure incorrectly favours the chosen model over the true model, when an incorrect model is chosen. Asterisk (*): mean $log_{10}(BF) \geq 1.30$, indicating strong relative support for the chosen model; dagger (†): mean $log_{10}(BF) \geq 2.18$, indicating very strong support (Kass and Raferty 1995).

A)

True model	*GeneralENM*	*Microsite*	*Multidimension*	*GrowCold*	*GrowDrought*	*LocalAdaptation*
	Model selected by ABC procedure					
GeneralENM	**52**	7	7	6	19	9
Microsite	6	**80**	4	6	1	3
Multidimension	0	1	**74**	23	1	1
GrowCold	0	1	25	**74**	0	0
GrowDrought	29	11	4	4	**47**	5
LocalAdaptation	3	3	0	2	2	**90**

B)

True model	*GeneralENM*	*Microsite*	*Multidimension*	*GrowCold*	*GrowDrought*	*LocalAdaptation*
	Model selected by ABC procedure					
GeneralENM	**0.26**					
Microsite		**2.12***				
Multidimension			**1.90***			
GrowCold				**1.38***		
GrowDrought					**0.69**	
LocalAdaptation						**5.00**†

C)	Model selected by ABC procedure					
True model	*GeneralENM*	*Microsite*	*Multidimension*	*GrowCold*	*GrowDrought*	*LocalAdaptation*
GeneralENM		0.48	0.44	0.68	0.32	0.59
Microsite	0.26		0.70	0.54	1.41*	0.29
Multidimension	-	0.38		0.43	1.00	0.48
GrowCold	-	0.62	0.37		-	-
GrowDrought	0.26	0.72	0.73	0.92		0.79
LocalAdaptation	0.33	0.27	-	0.72	0.80	

Figure 1. Geographic distribution of canyon live oak (grey shading; according to Little 1971) and sampling localities, where the size of the black circle corresponds to the number of individuals collected (sampling localities that are very close together were combined). Numbers on the black circles indicate populations as follows: (1) Peninsular Ranges, (2) Transverse Ranges, (3) Southern Coast Ranges, (4) Northern Coast Ranges and Klamath Mountains, (5) Southern Sierra Nevada, and (6) Northern Sierra Nevada. Several small, disjunct portions of the species distribution located east of the depicted range are not shown.

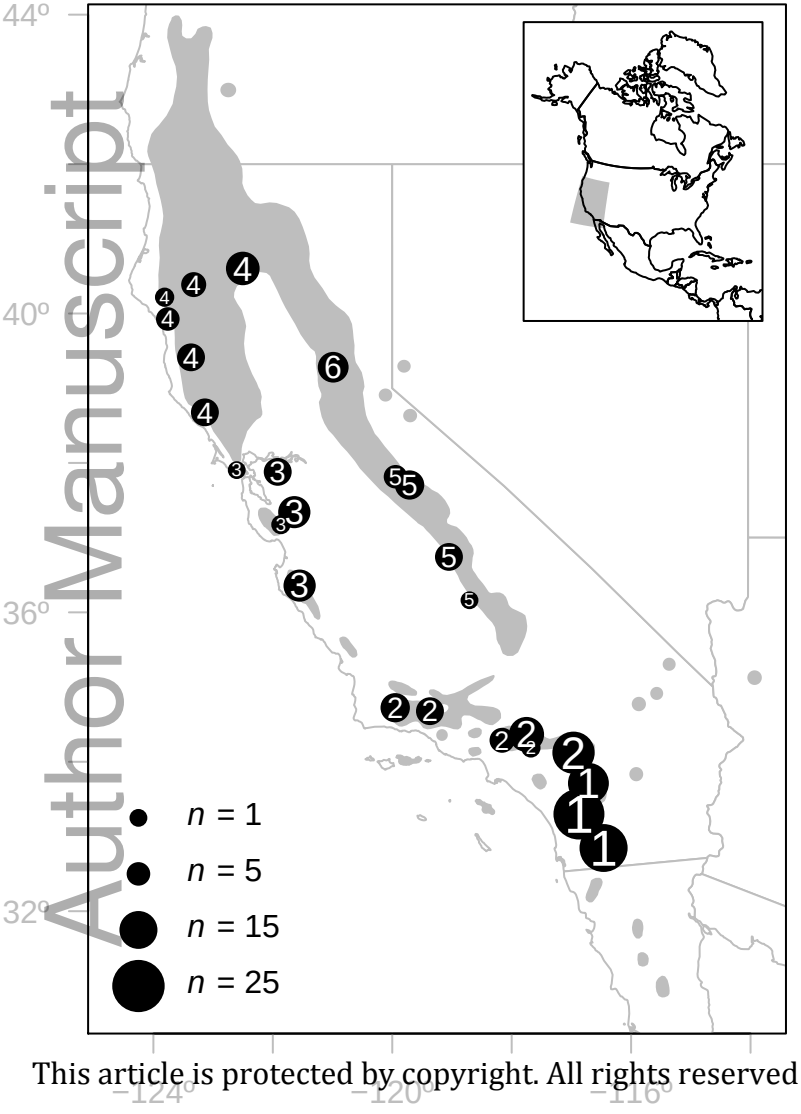
Figure 2. Dynamic ecological niche model used for demographic simulations, with an example illustrated for the *GeneralENM* model. Demes representing ancestral source populations (extracted from the areas of highest habitat suitability during the Last Glacial Maximum, LGM; see *Materials and Methods* for details) are initiated (grey arrow) within the LGM landscape at 21.5 ka. Demes are allowed to colonize the landscape, with carrying capacity and migration rate of each deme scaled relative to habitat suitability (coloured grid cells). Habitat suitability then shifts (black arrows) to that of intermediate and current time periods as the simulation progresses. One third of the total number of generations is simulated under each of the LGM, intermediate, and current landscapes.

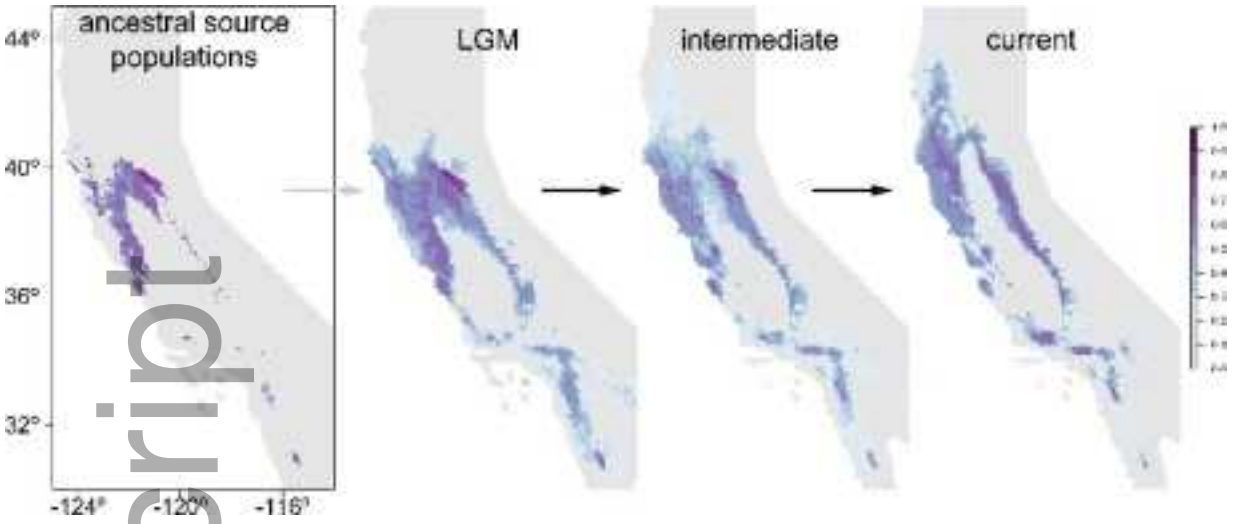
Figure 3. Habitat suitability for canyon live oak during the Last Glacial Maximum (21.5 ka) from ecological niche models constructed for each of the iDDC models.

Figure 4. Prior and posterior distributions of model parameters for the two most supported models, *GeneralENM* (a-d) and *GrowDrought* (e-h). Grey shading: prior distribution; dotted black line: posterior distribution before the ABC-GLM procedure; solid black line: final posterior distribution following ABC-GLM. K_{max}, carrying capacity; N_{anc}, ancestral population size; m, migration rate; μ, microsatellite mutation rate.

Figure 5. Distribution of posterior quantiles of true parameter values from 1,000 pseudo-observed datasets, used to assess bias in parameter estimation for the two

most supported models, the *GeneralENM* (a-d) and *GrowDrought* (e-h) models. Posterior quantiles (grey bars) are compared to a uniform distribution (dashed black line). The *p*-values test for deviation from a uniform distribution using a Kolmogorov-Smirnov test, with *p*-values less than 0.05 indicating bias in parameter estimation. K_{max}, carrying capacity; N_{anc}, ancestral population size; m, migration rate; μ, microsatellite mutation rate.





mec_13804_f2.tif